

IMPROVING THE ESP ACCURACY WITH COMBINATION OF PROBABILISTIC FORECASTS

Seung-Oh Yu¹ and Young-Oh Kim²

¹Assistant Research Associate, Research Institute of Engineering Science, Seoul National University

²Assistant Professor, School of Civil, Urban & Geosystem Engineering, Seoul National University

Abstract: Aggregating information by combining forecasts from two or more forecasting methods is an alternative to using forecasts from just a single method to improve forecast accuracy. This paper describes the development and use of a monthly inflow forecast model based on an optimal linear combination (OLC) of forecasts derived from naive, persistence, and Ensemble Streamflow Prediction (ESP) forecasts. Using the cross-validation technique, the OLC model made 1-month ahead probabilistic forecasts for the Chungju multi-purpose dam inflows for 15 years. For most of the verification months, the skill associated with the OLC forecast was superior to those drawn from the individual forecast techniques. Therefore this study demonstrates that OLC can improve the accuracy of the ESP forecast, especially during the dry season. This study also examined the value of the OLC forecasts in reservoir operations. Stochastic Dynamic Programming (SDP) derived the optimal operating policy for the Chungju multi-purpose dam operation and the derived policy was simulated using the 15-year observed inflows. The simulation results showed the SDP model that updated its probability from the new OLC forecast provided more efficient operation decisions than the conventional SDP model.

Keywords: ensemble streamflow prediction, optimal linear combination, stochastic dynamic programming

1. INTRODUCTION

A reliable streamflow forecast would be invaluable to water resource planners and managers but the mid- and long-term forecasts are seldom practically accurate in many countries. Korea is not an exception. In Korea, a simple categorical forecasting approach such as the 'Water Supply Outlook', which forecasts monthly streamflows at major control points for three categories, has been used: below normal flow, normal flow, and above normal flow. However, such forecasts have been insufficient for giving useful information to the water man-

agers and Ensemble Streamflow Prediction (ESP) technique could be a better alternative for the Water Supply Outlook (Kim et al., 2001). ESP is a well-known probabilistic forecasting technique of the National Weather Service, USA. Jeong and Kim (2002) applied ESP for probabilistic forecasting of the monthly inflows to the Chungju multipurpose dam in Korea. They performed an error analysis on their resulting probabilistic forecasts and found that the modeling error was dominant in winter and early spring (i.e. the dry season), while the meteorological error was dominant in summer (i.e. the flood season). Their conclusion suggested that

the rainfall-runoff model should be improved for the dry season while the reliable meteorological forecast should be utilized during the flood season. The latter conclusion has already been studied by Stedinger and Kim (2002), and thus, in our study, we addressed the former conclusion. Instead of improving the rainfall-runoff model itself, we attempted to improve the model outcome. Various methods and sources are available to produce forecasts, depending on the situation. To consider such multiple forecasts, one can select either the best forecast or aggregate multiple forecasts into a single forecast. The second option is called combined forecasting. Such combined forecasting is popular in economics, but there have only been a few attempts made in hydrologic forecasting (e.g. McLeod et al., 1987; Kim et al., 2003). Furthermore, combined probabilistic forecasting has been very rare even in economics. In developing a consensus probabilistic rainfall forecast in Australia, Casey (1995) proposed a statistical procedure based on an optimal linear combination of four probabilistic forecast models. In this study, we updated the flow probabilities every month after a new probabilistic forecast is calculated using the optimal linear combination. The main objective of this research is to improve the ESP accuracy for the Chungju multi-purpose dam and to assess the value of the improved forecasts for reservoir operations.

2. PROBABILISTIC FORECASTING

Hydrologic forecasting is classified into deterministic forecasting and probabilistic forecasting. Deterministic forecasting gives a specific value, whereas probabilistic forecasting suggests the likelihood of an event. Probabilistic forecasts range from 0 (event cannot occur) to 1.0 (event is certain to occur).

2.1 ESP Forecast

ESP runs a rainfall-runoff model with multiple meteorological inputs to generate an ensemble of possible streamflow hydrographs. A generated streamflow ensemble is a function of the initial hydrological states at the time of forecast in the rainfall-runoff model. The best probability density function (referred to as the ESP pdf hereafter) can then be fitted to the generated streamflow ensemble to describe the likelihood of an event occurring during a certain time period being forecasted. In this study, the entire range of streamflow (Q_t) was divided into three categories: low, medium, and high flow. A probability was then assigned to each category from the ESP pdf. The low, medium, and high flow probabilities (denoted \Pr^L , \Pr^M , and \Pr^H) were computed, respectively, using

$$\begin{aligned} \Pr^H &= 1 - \Pr[Q_t \leq q_t^U] \\ \Pr^M &= \Pr[Q_t \leq q_t^U] - \Pr[Q_t \leq q_t^L] \\ \Pr^L &= \Pr[Q_t \leq q_t^L] \end{aligned} \quad (1)$$

where q_t^L and q_t^U are the lower and upper limits of the medium flow category, respectively. In this study, the lower and upper limits are defined as the 33.3% and 66.7% cumulative quantiles, respectively, of a pdf fitted to the historical streamflow data.

2.2 Persistence Forecast

A probabilistic forecast using the persistence characteristic of streamflow may offer reasonable accuracy over a certain period. Such persistence between two adjacent forecasting times cannot be explicitly considered in ESP, although the initial condition of the ESP rainfall-runoff model may imply persistence. Therefore, the persistence forecast may compensate for this shortcoming of ESP when both forecasts are

combined.

This study employed a parametric approach that was often used to derive the transition probability in stochastic dynamic programming (e.g. Kim & Palmer, 1997). Assuming that two random variables follow a bivariate normal distribution, the conditional mean ($\mu_{Q_t|Q_{t-1}}^i$) and variance ($\sigma_{Q_t|Q_{t-1}}^2$) of Q_t given $Q_{t-1}=q_{t-1}^i$ for a bivariate case are, respectively,

$$\mu_{Q_t|Q_{t-1}}^i \equiv E(Q_t|Q_{t-1} = q_{t-1}^i) = \mu_{Q_t} + \rho_{Q_t, Q_{t-1}} \left(\frac{\sigma_{Q_t}}{\sigma_{Q_{t-1}}} \right) (q_{t-1}^i - \mu_{Q_{t-1}}) \quad (2)$$

$$\sigma_{Q_t|Q_{t-1}}^2 \equiv Var(Q_t|Q_{t-1} = q_{t-1}^i) = \sigma_{Q_t}^2 (1 - \rho_{Q_t, Q_{t-1}}^2) \quad (3)$$

where $\rho_{Q_t, Q_{t-1}}$ = the lag-1 correlation coefficient between Q_t and Q_{t-1} . Using the above mean and variance, the low, medium, and high flow probabilities, respectively are calculated as follows,

$$Pr^{Hi} = \int_{q_t^U} f_{Q_t|Q_{t-1}}(Q_t) dQ_t = 1 - \Phi \left[\frac{q_t^U - \mu_{Q_t|Q_{t-1}}^i}{\sigma_{Q_t|Q_{t-1}}} \right] \quad (4)$$

$$Pr^{Mid} = \int_{q_t^L}^{q_t^U} f_{Q_t|Q_{t-1}}(Q_t) dQ_t = \Phi \left[\frac{q_t^U - \mu_{Q_t|Q_{t-1}}^i}{\sigma_{Q_t|Q_{t-1}}} \right] - \Phi \left[\frac{q_t^L - \mu_{Q_t|Q_{t-1}}^i}{\sigma_{Q_t|Q_{t-1}}} \right] \quad (5)$$

$$Pr^{Lo} = \int_{-\infty}^{q_t^L} f_{Q_t|Q_{t-1}}(Q_t) dQ_t = \Phi \left[\frac{q_t^L - \mu_{Q_t|Q_{t-1}}^i}{\sigma_{Q_t|Q_{t-1}}} \right] \quad (6)$$

where $f_{Q_t|Q_{t-1}}$ = a normal pdf of Q_t on the condition of Q_{t-1} , and $\Phi[\cdot]$ = a standardized normal cdf.

2.3 Naive Forecast

If there is no deterministic forecasting technique available, only the average value of historical data can be used to make a forecast. This type of forecasting is called the deterministic naive forecast. For probabilistic forecasting, the naive forecast assigns a $1/J$ probability to each of J categories. In this study, each naive forecast has a 33.3% probability since three categories such as the low, medium, high flow are considered. A forecasting technique, therefore, shall be considered useful only if that technique is superior to the naive forecast. We expect that a technique combining forecasting techniques with the naive forecast prevents it from performing worse than the naive.

2.4 Optimal Linear Combination

No one method will produce the optimal forecast in all cases. An alternative approach is to combine the forecasts from two or more methods in accordance with their relative performances. By doing so, it is expected that the strengths of each method could be exploited (McLeod et al., 1987). A consensus forecast with high skill can be produced with an OLC (Piechota et al., 1998) of the three forecast models discussed above. The three forecast models such as the naive, persistence, and ESP forecasts, are combined into a final consensus forecast as follows,

$$Pr_{OLC} = a Pr_N + b Pr_E + c Pr_P \quad (7)$$

Determining an appropriate weight for each individual forecast is essential in the OLC approach. Appropriate weights are determined objectively by minimizing the error, which is measured using a forecast skill score such as the following Half Brier score (Brier and Allen,

1951),

$$HBS = \frac{1}{N} \sum_{j=1}^J \sum_{i=1}^N (\delta_{ij} - \phi_{ij})^2 \quad (8)$$

where ϕ_{ij} = the forecast probability that the event will occur in category j ; and δ_{ij} takes on a value of 1 or 0, depending on whether or not the event occurred in category j .

3. STOCHASTIC DYNAMIC PROGRAMMING

3.1 Formulation and Discretization

Reservoir operations always face uncertainties in inflow, water demand, and other stochastic variables. As a powerful stochastic optimization technique, SDP derives reservoir operating policies by solving the following recursive equation (Stedinger et al., 1984),

$$f_t(S_t, X_t) = \max_{R_t} E_{Q_t|X_t} \{ B_t(S_t, Q_t, R_t) + E_{X_{t+1}|Q_t, X_t} [f_{t+1}(S_{t+1}, X_{t+1})] \} \quad (9)$$

$$S_{t+1} = S_t + Q_t - R_t \quad (10)$$

$$R_t =$$

$$\min \{ \max [R_t^*, S_t + Q_t - S_{\max}], S_t + Q_t - S_{\min} \} \quad (11)$$

where S_t is the reservoir storage, Q_t is the inflow, X_t is the hydrologic state variable, S_{\min} is the minimum reservoir storage, S_{\max} is the maximum reservoir storage, R_t is the actual release, R_t^* is the target release derived from the optimal policy, $B_t(S_t, Q_t, R_t)$ is the immediate benefit function, and f_{t+1} is the optimal future value function at stage $t+1$.

If we ignore the hydrologic state variable, the SDP recursive equation, called SDP-N in this study, can be simply written as

$$f_t(S_t) = \max_{R_t} E_{Q_t} \{ B_t(S_t, Q_t, R_t) + f_{t+1}(S_{t+1}) \} \quad (12)$$

To solve the Equation 12, the state and stochastic variables such as S_t and Q_t are often represented with discrete values as shown in Equation 13,

$$f_t(S_t^k) = \max_{R_t} \sum_{i=1}^I \Pr(Q_t^i) \{ B_t(S_t^k, Q_t^i, R_t) + [f_{t+1}(S_{t+1}^i)] \} \quad (13)$$

where S_t^k is the k th discrete level of the reservoir storage state variable at the beginning of stage t , and Q_t^i is the i th discrete level of the stochastic inflow variable at stage t . The discretizing schemes for the state and stochastic variable will be described in the following section. In this study, the storage state S was discretized using the Savarenskiy's scheme (Klemes, 1977). The stochastic inflow, Q_t , was discretized into 7 states. For a standard normal distribute, Pegram et al. (1991) suggested -2.03, -1.18, -0.56, 0, 0.56, 1.18, and 2.03 for the discrete states and 0.054, 0.137, 0.198, 0.222, 0.198, 0.137, and 0.054 for the corresponding probabilities.

3.2 Solution Procedure and Re-Optimization

The SDP recursive equation is usually solved in a backward direction from the terminal stage for all possible combinations of the characteristic values. As mentioned earlier, the storage state variable is discretized into some representative values to allow numerical approximation of the future value function and the release policy. However, in actual operations, a calculated system state generally falls between these discrete values, requiring interpolation within the policy table.

To address this issue, Tejada-Guilbert et al.

Table 1. Half-Brier Scores of the 4 forecasting techniques

Month	1	2	3	4	5	6	7	8	9	10	11	12	Avg.
ESP	0.901	0.841	0.737	0.815	0.616	0.688	0.745	0.695	0.628	0.503	0.548	0.639	0.696
Pers.	0.563	0.436	0.524	0.618	0.548	0.664	0.634	0.672	0.664	0.737	0.622	0.482	0.597
Naive	0.667	0.667	0.667	0.667	0.667	0.667	0.667	0.667	0.667	0.667	0.667	0.667	0.667
OLC	0.599	0.436	0.535	0.629	0.570	0.712	0.664	0.711	0.675	0.518	0.601	0.525	0.598

*The best performance marked in bold.

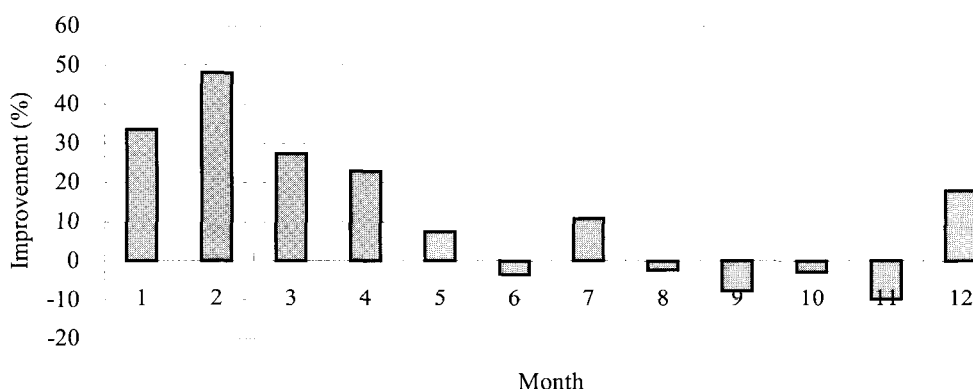


Figure 1. Percentage improvement of OLC against ESP

(1993) introduced the idea of reoptimization at each stage with the actual system storage, rather than interpolation within the policy table. Different from the general backward method in SDP, this reoptimization technique moves forward and calculates the optimal release. At each decision stage, a one-stage SDP optimization chooses an optimal release for the current state of the system to maximize the sum of the immediate benefit and the future value functions. In this study, the future value function that has been calculated from the backward SDP was used in the reoptimization procedure. Furthermore, the reoptimization procedure has another advantage, that is, discrete probabilities in SDP can be updated as a new forecast becomes available from OLC.

4. CASE STUDY

4.1 Study Basin

This study used monthly rainfall, inflow, and evaporation data of the Chungju dam for an uninterrupted 36-year period from 1966 to 2001. The Chungju dam basin consists of three sub-basins, each of which has several gauging stations. The observed water level was transformed to the discharges (i.e. inflows) using the stage-discharge curve. All records were obtained from the Korea Water Resources Corporation website at <http://wamis.kowaco.or.kr>.

4.2 The Probabilistic Forecasts

The ESP technique was applied to making 1-month ahead runoff forecasts from 1986 to

2001 when the observed runoff data are available. For each forecasting month, the historical rainfall and evaporation scenarios for the period from 1966 to 2001 were input to a rainfall-runoff model called TANK to generate the streamflow scenarios. Because the cross-validation technique was used in this study, we used the runoff data on the 35-year for calibration and the remaining year runoff data for the verification. Thus an ensemble of 35 runoff scenarios was generated at each forecasting month. To set the initial condition of each forecasting month, the TANK model was pre-ran for a two-month warming-up period. To issue a probabilistic forecast, the best probability density function was fit to the runoff ensemble each month. Through the goodness-of-fit test it was found that the 2-parameter lognormal distribution provided the best fit for most part of the verification period.

The persistence and the naive probabilistic forecasts were also carried out for the same period as was done by ESP. The forecast probability distributions of ESP, Persistence, and Naive were then discretized with the low, medium, and high flows to issue the probabilistic forecasts for each forecasting month.

As mentioned before, this study combined the ESP, Persistence, and Naive forecasts through the OLC approach. The weights were increased by 0.01 from 0 to 1 to search the best value set that produces the lowest Half-Brier score. Table 1 reports the Half-Brier scores of the four techniques: OLC improved the accuracy of ESP for the dry season but performed worse than Persistence. For the calibration period where the weights a , b , and c were estimated, OLC performed best, but this performance was not maintained for the verification period. This inconsistency problem should be explored further

in future research. Figure 1 summarizes the percentage improvement of OLC against ESP. On average, OLC achieved 11.83% in the HBS. As expected, the improvement was made in the dry season from December to May when the OLC technique was used.

4.3 Value of the OLC Forecasts in Reservoir Operations

In this thesis, two SDP models were used to obtain the optimal release policy for the Chungju multi-purpose dam: SDP-N and SDP-O. SDP-N is a classic SDP model moving backward while SDP-O is the model which can update every month its probability, which results from the OLC approach. Both models were reoptimized as described in the previous section. With such reoptimization procedure, SDP-O could update its OLC probability every month while SDP-N could update only the current storage since SDP-N used no forecast.

The primary objective for the monthly operations of the Chungju dam is to release water as demanded in the downstream area, followed by that needed for the hydropower generation. The objective function and constraints are as follows,

$$\max_{R_t} \sum_{t=1}^{12} (w_1 F_{1,t} - w_2 F_{2,t}) \quad (14)$$

where $F_{1,t}$ is the benefit from the hydropower generation at month t and $F_{2,t}$ is the loss from the water shortage at downstream at month t . The benefit rate (w_1) was set to 53.56 V/kWh while three values of the water shortage penalty (w_2) were considered in this study: 53.56 V/m^3 , 76.49 V/m^3 , and 122.35 V/m^3 .

The release constraint represents the amount of water that can be discharged from the

Table 2. Simulation results for 3 objective functions

Model		Annual Energy Product (GWh)	Annual Water Shortage (MCM)	Annual Spillway Release (MCM)	Annual Net Benefit ($\times 10^8$ W)
OBJ 1	SDP-O	790.2	127.9	830.5	354.7
	SDP-N	779.2	122.3	873.2	352.0
	Actual	863.0	637.6	14.3	120.7
OBJ 2	SDP-O	778.4	114.2	902.1	354.7
	SDP-N	752.9	104.5	1070.0	352.0
	Actual	863.0	637.6	14.3	120.7
OBJ 3	SDP-O	772.3	100.9	996.5	354.7
	SDP-N	748.6	94.1	1108.8	352.0
	Actual	863.0	637.6	14.3	120.7

*OBJ 1: $w_1 = 53.56 \text{ V/k}$, $w_2 = 53.56 \text{ V/m}^3$

OBJ 2: $w_1 = 53.56 \text{ V/k}$, $w_2 = 76.49 \text{ V/m}^3$

OBJ 3: $w_1 = 53.56 \text{ V/k}$, $w_2 = 122.35 \text{ V/m}^3$

Chungju dam, ranging from the instreamflow requirement (10.6 CMS) to the turbine capacity (1980 CMS). The storage state variable was discretized into 30 units by using the Savarenkiy's method and the stochastic variable into 7 units as mentioned before. For each discretized state, the optimal solution was searched by increasing the decision variable (i.e. release) by 28.5 CMS. Moving backward, the iteration of the SDP recursive equation stopped when the difference between two consecutive iterations was smaller than 0.01 for all the discrete state and months.

The release policies derived with both SDP models were simulated using the observed monthly inflow data from 1987 to 2001. Using again the cross validation scheme, the simulation procedure started in January of 1987 when the actual initial storage data began to become available. We objectively compared the SDP simulation results with those using the actual release data with respect to three operational aspects, the hydropower generation, the water shortage, and the spill. The simulation results are given in Table 2.

The simulation results showed that the hydropower generation and the water shortages of both SDP-N and SDP-O decrease as the penalty on the water shortage increase. With respect to the spill and the hydropower generation, the SDP models did not performed better than the actual operations. As for the three objective functions, the spill calculated from SDP-O was less than that from SDP-N, but it was much larger than the actual. Because this study did not consider the minimization of the spill in the SDP objective function, so that the proposed SDP models spilled water in the flood season, i.e., July, August, and September. For the dry season, almost zero spill occurred during the irrigation period. Since the actual operation released considerably more water to mitigate the flood risk during the flood season, it generated more hydropower energy than SDP-N and SDP-O during that season and consequently throughout the year.

The annual net benefit of SDP-O is greatest for all of three objective functions. As compared with the result of SDP-N, SDP-O showed improvements of 0.8 %, 1.9 %, and 1.5 %. These

improvements correspond to 2.7 hundred million, 6.3 hundred million, and 4.4 hundred million won, respectively.

5. CONCLUSIONS

To improve the ESP forecasting accuracy, two additional probabilistic forecasts such as the persistence and the naive forecasts were combined with the ESP forecast using OLC that minimizes the HBS. The proposed methodology was applied to the probabilistic forecasting of 1-month ahead inflows to the Chungju dam. The forecasting results for the verification period associated with the cross-validation showed OLC improved ESP especially during the dry season. On average, OLC achieved 11.83% improvements over ESP in terms of HBS. The forecasts shown here can be made in terms of the conditional probability of exceedance that can be helpful for managing reservoirs, forecasting future allocations, and establishing drought management plans.

To examine the value of the improved forecasts in reservoir operations, two versions of stochastic dynamic programming models such as SDP-N and SDP-O were developed. The SDP-derived release policies were simulated with the observed inflow data. The greater penalty causes the smaller power generation. With respect to the annual net benefit, SDP-O shows the better performance than the actual release case as well as SDP-N.

To conclude, the use of the OLC probabilistic forecasts with SDP can provide more efficient operating decisions than SDP without forecasts due to the improved accuracy of OLC over ESP.

ACKNOWLEDGEMENT

This research was supported by a grant (code 1-6-1) from Sustainable Water Resources Research

Center of 21st Century Frontier Research Program and was also supported by the Brain Korea 21.

REFERENCES

- Brier, G. W., and Allen, R. A. (1951). "Compendium of meteorology." *Am. Meteorol. Soc.*, 841-848.
- Casey, T. (1995). "Optimal linear combination of seasonal forecasts." *Aust. Meteorol. Mag.*, 44, 219-224.
- Faber, B. A., and Stedinger, J. R. (2001). "Reservoir optimization using sampling SDP with ensemble streamflow prediction (ESP) forecasts." *J. Hydro.*, 249(1-4), 113-133.
- Jeong, D. I., and Kim, Y.-O. (2002). "Forecasting monthly inflow to Chungju dam using ensemble streamflow prediction." *J. Korean Society of Civil Engineers*, 22(3-B), 321-331 (in Korean).
- Kim, Y.-O., Jeong, D. I., and Kim, H. S. (2001). "Improving water supply outlook in Korea with ensemble streamflow prediction." *Water International*, 26(4), 563-568.
- Kim, Y.-O., Jeong, D.I., and Ko, I. H. (2003). "Combining rainfall-runoff models for improving ensemble streamflow prediction." *J. Hydro.* (submitted).
- Kim, Y.-O., and Palmer, R. N. (1997). "Value of seasonal forecasts in bayesian stochastic programming." *J. Water Resour. Plng. and Mgmt.*, 123(6), 327-335.
- Klemes, V. (1977). "Discrete representation of storage for stochastic reservoir optimization." *Water Resour. Res.*, 13(1), 149-158.
- McLeod, A. I., Noakes, D. J., Hipel, K. W., and Thompstone, R. M. (1987). "Combining hydrologic forecasts." *J. Water Resour. Plng. and Mgmt.*, 113(1), 29-41.
- Pegram, G., Stedinger, J., Payne, K., and Born, P. (1991). "Statistical analysis program for

- SDP optimization.” *Cornell University*, Ithaca, New York, 1-27 (unpublished manuscript).
- Piechota, T. C., Chiew, F. H. S., Dracup, J. A., and McMahon, T. A. (1998). “Seasonal streamflow forecasting in eastern Australia and the El Niño-Southern Oscillation.” *Water Resour. Res.*, 34(11), 3035-3044.
- Stedinger, J. R., and Kim, Y.-O. (2002). “Updating ensemble probabilities based on climate forecasts.” *Conf. on Water Resour. Plng. and Mgmt.*, ASCE, Roanoke, VA, Session C2(CD-ROM).
- Stedinger, J. R., Sule, B. F., and Loucks, D. P. (1984). “Stochastic dynamic programming models for reservoir operation optimization.” *Water Resour. Res.*, 20(11), 1499- 1505.
- Tejada-Guilbert, J. A., Johnson, S. A., and Stedinger, J. R. (1993). “Comparison of two approaches for implementing multireservoir operating policies derived using stochastic dynamic programming.” *Water Resour. Res.*, 29(12), 3969-3980.
-
- ¹Assistant Research Associate, Research Institute of Engineering Science, Seoul National University, San 56-1, Shillim-dong, Gwanak-gu, Seoul, 151-742, Korea.
(E-mail: skull1@snu.ac.kr)
- ² Corresponding Author, Assistant Professor, School of Civil, Urban & Geosystem Engineering, Seoul National University, San 56-1, Shillim-dong, Gwanak-gu, Seoul, 151-742, Korea.
(E-mail: yokim05@snu.ac.kr)